This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problems Mailbox.

Office de la propriété intellectuelle du Canada

Un organisme d'Industrie Canada

Canadian Intellectual Property Office

An Agency of Industry Canada

Bureau canadien des brevets

Certification

Canadian Patent Office

Certification

La présente atteste que les documents ci-joints, dont la liste figure ci-dessous, sont des copies authentiques des documents déposés au Bureau des brevets.

This is to certify that the documents attached hereto and identified below are true copies of the documents on file in the Patent Office.

Specification and Drawings, as originally filed, with Application for Patent Serial No: **2,301,394,** on March 20, 2000, by **SPACEBRIDGE NETWORKS CORPORATION,** assignee of Eko Adi Wibowo, for "Method and System for Resource Allocation in Broadband Wireless Networks".

RECEIVED
JUN 1 1 2001
Technology Center 2600

S. Gregoire

Agent certificateur/Certifying Officer

March 23, 2001

Date

Canada

OPIC CIPO

ABSTRACT

This invention is primarily concerned with the method of resource allocation in broadband wireless networks that is executed by a resource allocation server typically located at hub of the network. More specifically, the invention is concerned with the method of distributing transmission bandwidth to requesting wireless access devices using VBDC and FCA mechanisms.

METHOD AND SYSTEM FOR RESOURCE ALLOCATION IN BROADBAND WIRELESS NETWORKS

FIELD OF THE INVENTION

The present invention generally relates to a method and system for allocating resources in broadband wireless networks. More specifically, the present invention relates to a method and system for allocating bandwidth in a fair and efficient manner to several users in a broadband wireless network

BACKGROUND OF THE INVENTION

A multiple access method for wireless networks can be viewed as being composed of two components, namely the primary and the secondary multiple access methods. The primary multiple access method is concerned with the general utilization of transmission bandwidth by wireless access devices. The secondary multiple access method is concerned with the scheduling of transmission of each wireless access device in the network.

The primary multiple access method for broadband wireless networks should be selected such that it allows for the most efficient use of bandwidth and most cost effective implementation, and is capable to support variable transmission rate as well as accommodate large number of users. There are three basic primary access methods for wireless communications, namely frequency division multiple access (FDMA), time division multiple access (TDMA) and code division multiple access (CDMA). In addition, there are a number of variants derived from the combination of these three basic access methods.

In FDMA, transmission bandwidth is divided into frequency bands or channels. A wireless access device has its own channel(s) to transmit its payloads, hence collision can be avoided. Another advantage of FDMA is inexpensive implementation, due to its relative simplicity compared to TDMA and CDMA technologies. Nevertheless, FDMA also has several disadvantages. First, the number of wireless access devices that the network can accommodate is limited to the number of channels. Second, sufficient guard bands are required to avoid interference between channels, which translates into some loss of efficiency in bandwidth utilization. In addition, there is little flexibility in handling traffic variability.

In TDMA, transmission bandwidth is divided into time slots. In the most simplistic case,

a wireless access device is assigned one or more time slots on a fixed periodic basis. In a more advanced case, a wireless access device is allowed to vary its bandwidth requirement on an as-needed basis, resulting in a bandwidth-on-demand scheme. As a result, a large number of wireless access devices can be supported by the network. Furthermore, TDMA technology provides flexibility in handling variability of network traffic. There are, however, problems associated with it. First, in order to achieve higher transmission speed, larger antenna size is required. This negatively impacts the cost of wireless access devices. Second, there is a need for accurate time synchronization, which results in additional hardware complexity. Third, there is still a need for guard bands (in terms of time), to give sufficient rooms to tolerate variable propagation and multi-path delays [1]. Nevertheless, TDMA technology is considered to be a good match for integrating multi-media services, such as voice and data, in the same network [2].

In CDMA, the whole transmission bandwidth is available to all wireless access devices at all time. The concept is in contrast to the previous two access mechanisms, where a wireless access device is given only a portion of the transmission bandwidth. In CDMA, signals from one wireless access device are allowed to overlap with those from other wireless access devices. Hence, multiple access capability is built-in, and, due to its random access nature, adding users to the network is a simple process. Each wireless access device is assigned a unique code. Coded signals can be extracted in the receiver side by using a complementary code. Each code is designed to be orthogonal to any other code, that is, each code is made to be as different as any other [3]. As a result, signal interference is minimal. Other advantages of CDMA technology include built-in security mechanism and robustness to various impairments in wireless channels or systems, such as multi-path fading, frequency selectivity and synchronization errors [4]. The major problem associated with CDMA technology is its low bit rate, which is primarily due to limitations in hardware implementation.

As none of these three primary access mechanisms possess all the qualities of an excellent access mechanism, researchers attempt to combine them in order to gain as many advantages and eliminate as many disadvantages as possible. An example of such combinations is Multi-Frequency TDMA (MF-TDMA).

In MF-TDMA, transmission bandwidth is divided into both frequency and time slots. A wireless access device is allowed to transmit at any frequency at any time slot, but typically not at more than one frequency simultaneously. This imposes a limit on the rate at which a wireless

access device can transmit, with positive impacts on antenna size and transmission power requirements. This invention is based on the use of MF-TDMA as the primary multiple access method.

MF-TDMA is a frame-based method. This means that time slots are organized into a regular structure. Slots in a frame are typically grouped according to their functions. These functions relate to how data transfer process is conducted in a MF-TDMA wireless network. Before a wireless access device is allowed to use any of the network resources, it has to register itself to the network. Next, it has to constantly synchronize itself to the network timing so that it does not disrupt or interfere with any other wireless access devices. Then, when demand-based multiple access is employed, a wireless access device has to signal its request for uplink bandwidth to the process in charge of uplink bandwidth allocation, the resource allocation server. The wireless access device may transmit its data once its assignment message arrives. Based on this transmission procedure, time slots can be classified into five main function groups, namely log slots, acquisition slots, synchronization slots, request slots and data slots.

Log slots are used in the registration or log-in process, and are usually randomly accessed by wireless access devices. Acquisition slots are used to achieve synchronization prior to any data transfer. Synchronization slots are used by access devices to maintain synchronization to network timing. Request slots carry bandwidth request information from each access device, and consequently, are only useful for demand-based multiple access schemes. Finally, data slots are used to transfer actual network traffic.

The utilization of data and request slots is governed by the secondary multiple access method. The combined free-demand assignment multiple access method (CF-DAMA), on which the invention is based, was first introduced in [5]. CF-DAMA aims to maximize transmission bandwidth utilization by requiring each wireless access device to reserve or request bandwidth for its traffic. The resource allocation server allocates bandwidth in response to the access device's demand. Furthermore, unreserved data slots are freely distributed by the resource allocation server to wireless access devices. Advantages of the method are many, including efficient bandwidth utilization, robustness, controlled fairness and stable operation up to very high traffic loads

To date, several flavours of CF-DAMA have been conceived. In [6] a version of CF-DAMA with piggy-back request slots was described. In this version of CF-DAMA, a wireless

access device with no requests queued in the resource allocation server's request queue has to wait until it receives free assignment of a data slot to send its traffic and piggy-back its request in the header section of the data packet. Since this may create unfavourable conditions for wireless access devices with no queued requests at high traffic load situations, [6] suggested that such access devices be given a higher priority in receiving free assignment, that is, all wireless access devices with no queued requests will be considered for free assignment before those with queued requests. This method possesses all of the advantages associated with CF-DAMA described earlier. This method, however, does not support multi-priority traffic. As a result, it cannot be used to provide integrated voice and data services.

In [7] a version of CF-DAMA with fixed-assigned request slots was described. In this method, a separate signaling channel was provided, in which each wireless access device has one reserved request slot for its use. This approach solves the potential unfairness problem associated with the piggy-back version at a cost of more signaling overhead and limited number of access devices that the network can accommodate. Furthermore, the authors suggested to integrate voice and data applications by adding implicit reservation for voice traffic, in the sense that an assigned slot for a voice application will be implicitly kept in subsequent frames until the reserving access device transmits a release request. The method, however, cannot be extended to handle multiple classes of service and maintain quality of service guarantees.

In [8] a version of CF-DAMA with four request/assignment strategies was introduced. These strategies are constant rate allocation (CRA), rate-based dynamic capacity (RBDC) allocation, volume-based dynamic capacity (VBDC) allocation and free capacity assignment (FCA). In CRA, the resource allocation server provides a wireless access device with a number of data slots per frame for each and every frame while required. This strategy is aimed toward real-time services, i.e., those that have strict constraints on delay and delay variation. Next, in RBDC, the access device may request a variable number of data slots at each frame. The resource allocation server, in turn, provides a wireless access device with a guarantee on the maximum number of data slots it can make available at each frame if requested by the access device. This strategy is intended toward services that can tolerate delay. Furthermore, in VBDC, the access device signals its requests in terms of total number of data slots (i.e., volume). The resource allocation server, however, does not provide any guarantee on the availability of data slots. This strategy is intended toward services that can tolerate delay and delay variation. Finally, the

resource allocation server attempts to maximize bandwidth utilization by distributing unassigned data slots to all eligible access devices using FCA. These request/assignment strategies can be flexibly combined to match quality-of-service (QoS) requirements of a given class of service. The mapping of these strategies to ATM service classes was also illustrated in [8].

The resource allocation method proposed by [8] is composed of two stages. In the first stage, the resource allocation server calculates the number of data slots reserved through CRA and RBDC and requested through VBDC. In the second stage, the resource allocation server assigns to each access device the guaranteed number of data slots (i.e., CRA and the requested RBDC up to the maximum RBDC value). The resource allocation server also attempts to satisfy the access device's VBDC request in full (the resource allocation server stores in memory each VBDC request from the access device that has not been granted in full). Finally, similar to [6], the resource allocation server gives a higher priority to access devices that have no bandwidth assignment in a frame (through CRA, RBDC or VBDC) to receive FCA than those with one.

The request/assignment and resource allocation strategies proposed by [8] are capable of handling multiple classes of service and maintaining QoS guarantees associated with each class of service. Nevertheless, the attempt to satisfy each access device's VBDC request in full brings two shortcomings. First, this approach increases in short-term traffic burstiness, which can potentially increase congestion loss. Second, this approach also increases data transfer delay at high traffic loads. By attempting to satisfy each access device's VBDC request in full, the number of access devices that may receive bandwidth allocation through VBDC in a frame is reduced. As a result, at high traffic loads, it will take longer for an access device to receive bandwidth allocation through VBDC or to send VBDC requests. Hence, data transfer delay increases.

A good method to tackle these two shortcomings was described in [9]. In [9] the resource allocation server uses tokens to determine the number of data slots it should make available to a given access device. The resource allocation server grants a token to a given access device periodically (the period of the token grant is inversely proportional to the service rate requested by the access device). The access device accumulates tokens it receives from the resource allocation server. Then, when the access device requests for bandwidth, the resource allocation server only attempts to provide a number of data slots up to the number of tokens that the access device has accumulated. This method is intended to reduce traffic burstiness and congestion loss.

This method should also reduce data transfer delay at high traffic loads, as more access devices will have bandwidth allocation in a frame. The method, however, does not attempt to maximize bandwidth utilization. The resource allocation server does not attempt to satisfy as many bandwidth requests as possible, even if bandwidth is available after the resource allocation server has exhausted all accumulated tokens in all access devices. In addition, the method does not consider granting free data slots through FCA. The invention addresses these shortcomings by using this available bandwidth to service any pending VBDC requests and to provide free capacity assignment.

An alternative method to address the two shortcomings of [8] was described in [10]. In [10] the resource allocation server may grant to each access device a fair share of best-effort bandwidth, which is bandwidth that is available after all bandwidth reservation through static rate (which is similar to CRA) and booked rate (which is similar to RBDC) has been satisfied. The resource allocation server will only consider granting this best-effort share to an access device if the access device requests more data slots than its booked rate. The strength of this method is its explicit support for service fairness. A fair-share of best-effort bandwidth is distributed fairly to access devices with pending requests. The shortcoming of this method is that by attempting to fairly distribute best-effort bandwidth to access devices, more lengthy operations are required to maximize bandwidth utilization in situations where the number of pending requests in each access device varies greatly.

Prior art references include:

1. Andrew J. Viterbi. A Perspective on the Evolution of Multiple Access Satellite Communication. IEEE Journal on Selected Areas in Communications, Vol. 10, No. 6, August 1992.
2. Shuzo Kato, et. Al. A TDMA Satellite Communication System for ISDN Services. IEEE Journal on Selected Areas in Communications, Vol. 10, No. 2, February 1992.
3. Uyless Black. Mobile and Wireless Networks. Prentice Hall, Inc. Upper Saddle River, 1996.
4. L. Schilling, et. Al. Broadband CDMA for Personal Communication Systems. IEEE Communications Magazine, November 1991.
5. Jahangir I. Mohamed. Combined Free/Demand Assignment Multiple Access Protocols for Packet Satellite Communications. Master Thesis. Concordia University, June 1993.
6. Jahangir I. Mohammed and Tho Le-Ngoc. Performance Analysis of Combined Free/Demand

Assignment Multiple Access (CFDAMA) Protocol for Packet Satellite Communications. Proceedings of the ICC'94 Conference, 1994, pp. 869-873.

7. S. V. Krishnamurthy and Tho Le-Ngoc. Performance of CF-DAMA Protocol with Pre-Assigned Request Slots in Integrated Voice/Data Satellite Communications. Proceedings of the ICC'95 Conference, 1995, pp. 1572-1576.
8. E. A. Wibowo, A. Iuoras, P. Takats, J. Lambadaris, M. Devetsikiotis. Guaranteeing Quality of Service in Packet-Switched Satellites by Medium Access Control. Proceedings of the CCBR'98 Conference, Ottawa, June 21st - 24th, 1998.
9. Jun Huang, Tho Le-Ngoc and Jeremiah F. Hayes. Broadband Satcom System for Multimedia Services. Proceedings of the ICC'96 Conference, 1996, pp. 906-910.
10. Guray Açar and Catherine Rosenberg. Algorithms to Compute Bandwidth on Demand Requests in a Satellite Access Unit. Proceedings of the Fifth Ka-Band Utilization Conference, Taormina, Sicily Island, Italy, October 18 - 20, 1999.
11. Shuang Deng, "Empirical Model of WWW Document Arrivals at Access Link", Conference Record ICC'96, 1996, pp. 1797-1802.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a method and system for allocating bandwidth in a fair and efficient manner to several users in a broadband wireless network.

The present invention provides a method for resource allocation applicable to a base station unit in broadband terrestrial wireless and satellite networks, comprising the steps of:

(i) performing a first scan of an access device table by calculating the number of data slots for prioritized VBDC grant and for normal VBDC grant to each wireless access device, and calculating the total number of data slots for prioritized VBDC grant and for normal VBDC grant to all wireless access devices in the access device table.

(ii) performing a second scan of the access device table by calculating the total number of data slots available for prioritized VBDC grant, for normal VBDC grant and for FCA grant, calculating the number of data slots to be allocated to each access device through prioritized VBDC grant, calculating the total number of credits available to each access device for prioritized VBDC grant, calculating the number of data slots to be allocated to each access device through normal VBDC grant, calculating the total number of data slots to be allocated to each access device, which is the sum of data slots reserved through CRA and RBDC as well as data slots obtained through prioritized and normal VBDC grants, determining if each access device is eligible to receive FCA, calculating the number of data slots to be allocated to each access device through FCA grant, and calculating the total number of credits available to each access device for FCA grant.

BRIEF DESCRIPTION OF THE DRAWINGS

Exemplary embodiments of the present invention will now be described, with reference to the attached drawings, in which:

Figure 1 is a simple Bandwidth-on-Demand network

Figure 2 is a flowchart of the resource allocation method – Stage 1

Figure 3 is a flowchart of the resource allocation method – Stage 2

Figure 4 is a flowchart of the resource allocation method – Stage 2 (Continued)

Figure 5 is a block diagram of a high-level software architecture for the implementation of the resource allocation method

Figure 6 is a diagram illustrating a sample simulation network model with a one hub

node and 8 wireless access devices

Figure 7 is a graph illustrating the average transfer delay for high-priority data Packets

Figure 8 is a graph illustrating an average transfer delay for low-priority data Packets

DETAILED DESCRIPTION

The resource allocation server is responsible for distributing uplink bandwidth based on requests from wireless access devices. In distributing uplink bandwidth, the resource allocation server must guarantee that it can satisfy each access device's CRA. In addition, the resource allocation server must also guarantee that it can satisfy each access device's RBDC request up to the maximum RBDC value. The resource allocation server is free to distribute using VBDC and FCA any uplink bandwidth unassigned after servicing CRA and RBDC requests. The resource allocation server, however, should be fair in distributing this unassigned bandwidth as service fairness brings about many advantages, including:

- Reductions in request access delays.
- Reductions in cell transfer delays.
- Increased traffic shaping effect.

A resource allocation method to fairly allocate unassigned uplink bandwidth to wireless access devices using VBDC and FCA is described below.

The novel resource allocation method is one component of the set of Media Access Control (MAC) protocols and algorithms of a Bandwidth-on-Demand (BoD) network. In order to understand the functionality and appreciate the importance of a resource allocation method, a high-level knowledge of how such a network operates is beneficial.

A simple BoD network may consist of a number of access devices and a base station (or alternatively, a hub) that is located at the center of the network (see Figure 1). These access devices are typically sharing uplink transmission bandwidth to the base station. Therefore, in order to reduce the occurrences of or avoid collisions, the base station typically controls access to the transmission bandwidth. Furthermore, in order to achieve high bandwidth utilization, it is advantageous if the base station is capable of granting access to the transmission bandwidth only to devices that need to send data to the base station (the term bandwidth-on-demand is hence used to describe such capability). For the base station to regulate access and provide bandwidth on demand to access devices, the following operations are then necessary:

- Each access device computes the amount of transmission bandwidth required to satisfy its traffic requirement.
- Each access device signals or requests this calculated amount of transmission bandwidth to the base station.
- The base station (more specifically, the resource allocation server that is located in the base station), using the resource allocation method, calculates the amount of transmission bandwidth it can grant to each access device based on reservations and/or requests from the corresponding access device.
- The base station builds and signals to all access devices an assignment table (called burst-time plan or BTP) containing information on the frequencies and times each access device that is granted access can send its data.
- Each access device scans each assignment table for information on the frequencies and times it can send its data. Each access device transmits its data at the specified frequencies and times.

Central to the above operations is the execution of the resource allocation method. Indeed, the resource allocation method significantly influences the attributes of the network, such as maximum sustainable throughput, data transfer delay and fairness.

The resource allocation method supports all four request/assignment strategies described in [8], namely CRA, RBDC, VBDC and FCA. The novelty of the resource allocation method lies in the treatment of VBDC requests and the distribution of FCA. Unlike the method described in [8], the resource allocation method does not maintain each pending VBDC request from each access device. Instead, the method accumulates all pending VBDC request from each access device. Furthermore, the method does not provide two priorities of VBDC requests as in [8]. Instead, the method treats the total VBDC request of a given access device as having two components, the prioritized VBDC component and the normal VBDC component. Finally, the method uses a more sophisticated method of providing FCA to access devices than that proposed in [8].

Each access device may be given a number of data-slot credits (or tokens) by the resource allocation server in every frame. The number of credits the resource allocation server grants in every frame to an access device is typically equivalent to the service rate requested by the access device. These credits are accumulated and form the prioritized VBDC component of the

corresponding access device. The resource allocation server uses an access device's accumulated credits to determine the access device's prioritized VBDC assignment.

The normal VBDC component of a given access device is equivalent to the number of data slots requested by the access device through VBDC less the access device's prioritized VBDC component. The resource allocation server will only attempt to satisfy an access device's normal VBDC component when bandwidth is available after the resource allocation server satisfies CRA, RBDC and prioritized VBDC components of all access devices.

Finally, if bandwidth is available after the resource allocation server satisfies CRA, RBDC, prioritized VBDC and normal VBDC components of all access devices, the resource allocation server attempts to distribute unassigned bandwidth through FCA. Similar to the approach for providing prioritized VBDC component, each access device may be given a number of data-slot credits by the resource allocation server in every frame for FCA. The resource allocation server uses an access device's accumulated FCA credits to determine the number of free data slots it can grant to the corresponding access device. Additionally, the resource allocation server applies a threshold to the number of data slots that will be assigned to the access device due to its CRA, RBDC, prioritized VBDC and normal VBDC components, above which the resource allocation server will not consider granting free slots to the corresponding access device.

Similar to the method presented in [8], the resource allocation server maintains an access device table, which contains information of each wireless access device registered in the network. The table is arranged in the form of a circular linked-list of access device entry. Each entry in the table contains information on resource requirement of a specific wireless access device. Furthermore, similar to the method presented in [8], the resource allocation method is composed of two stages. In the first stage, the resource allocation server scans the table and calculates the total requested bandwidth for CRA, RBDC and VBDC. In the second stage, the resource allocation server re-scans the table and assigns bandwidth to wireless access devices. Finally, similar to the method presented in [8], the resource allocation server performs these table scans in a round-robin fashion, visiting each access device entry one by one starting from the first device entry in the table. The resource allocation server may tag the first device entry that cannot be given its prioritized VBDC component in full or partially. Similarly, the resource allocation server may also tag one device entry for unmet normal VBDC component and one device entry

for unmet FCA component. The first device entry with the highest-priority unmet component[1] is made the first entry of the table, so that it can receive better service in the next execution of the resource allocation method.

As mentioned earlier, each entry in the table contains information on resource requirement of a specific wireless access device. With respect to VBDC and FCA, this information is as follows:

- Requested VBDC value (Dev.VBDC.ReqValue).

This information indicates the total number of data slots the wireless access device has requested through VBDC. This information is updated every time the resource allocation server receives a VBDC request from the access device, in which case the new VBDC request value is added to previous requests.

- Prioritized VBDC credit (Dev.VBDC.Credit).

This information indicates the number of data slots (or alternatively, tokens) the wireless access device may accumulate for VBDC at every frame due to unused VBDC credit. Prioritized VBDC credit can be a fractional value if fine token granularity is desired.

- Prioritized VBDC total credit (Dev.VBDC.TotalCredit).

This information indicates the total number of data slots the wireless access device has accumulated for VBDC up to the current frame due to unused VBDC credit.

- Prioritized VBDC maximum total credit (Dev.VBDC.MaxCredit).

This information indicates the maximum total number of data slots the wireless access device can accumulate for VBDC due to unused VBDC credit. This number is typically associated with the maximum size of data that can be transmitted at the peak rate.

- Current Prioritized VBDC value (Dev.VBDC.PrioValue).

This information indicates the number of data slots the wireless access device may have prioritized access to through VBDC for the considered frame.

- Current VBDC value (Dev.VBDC.Value).

This information indicates the number of data slots the wireless access device may have access to through VBDC for the considered frame.

- FCA threshold (Dev.FCA.Threshold).

---

1 The priority ordering, from highest to lowest, is as follows: prioritized VBDC component,

This information indicates the threshold on the number of data slots assigned to a wireless access device, above which the access device will not be considered for FCA.

- FCA credit (Dev.FCA.Credit).

This information indicates the number of data slots the wireless access device may accumulate for FCA at every frame due to unused FCA credit.

- FCA total credit (Dev.FCA.TotalCredit).

This information indicates the total number of data slots the wireless access device has accumulated for FCA up to the current frame due to unused FCA credit.

- FCA maximum total credit (Dev.FCA.MaxCredit).

This information indicates the maximum total number of data slots the wireless access device may accumulate for FCA due to unused FCA credit.

- Current FCA value (Dev.FCA.Value).

This information indicates the number of data slots the wireless access device may have access to through FCA for the considered frame.

- Current bandwidth assignment value (Dev.BandwidthAssign).

This information indicates the number of data slots the resource allocation server will allocate to the wireless access device for the considered frame through CRA, RBDC, VBDC and FCA.

Figure 2 shows the flowchart of stage 1 of the resource allocation method. As can be seen from the flowchart the first stage begins with the resource allocation server initializing the total CRA, total RBDC, total prioritized VBDC and total VBDC counters to zero. The resource allocation server then begins examining bandwidth reserved and requested by each wireless access device, starting from the first access device in the table.

The resource allocation server first sums the total bandwidth reserved through CRA and RBDC. The resource allocation server continues to examine each wireless access device entry if it has VBDC requests for the considered frame. The resource allocation server calculates two types of VBDC grants to an access device, namely prioritized VBDC grant and normal VBDC grant. The purpose of this prioritized grant is to shape traffic and provide fair access to all registered access devices in the network by distributing VBDC grants to as many access devices as possible.

---

normal VBDC component, FCA component.

The number of data slots that the resource allocation server may give prioritized access to the access device through VBDC is the minimum of the number of data slots per frequency channel less the number of data slots to be allocated through CRA and RBDC (ChannelCap – Dev.CRA.Value – Dev.RBDC.Value)2, the number of data slots that the access device has prioritized access to (Dev.VBDC.Credit + Dev.VBDC.TotalCredit, rounded down to the nearest integer value) and the number of data slots requested through VBDC (Dev.VBDC.ReqValue). The number of data slots that the resource allocation server may grant to the access device through VBDC is the minimum of the number of data slots per frequency channel less the number of data slots to be allocated through CRA and RBDC (ChannelCap – Dev.CRA.Value – Dev.RBDC.Value) and the number of data slots requested through VBDC (Dev.VBDC.ReqValue). The resource allocation server then updates the total number of data slots for prioritized VBDC allocation and the total number of data slots for VBDC allocation.

Figure 3 and Figure 4 show the flow chart of stage 1 of the resource allocation method. As can be seen from the flow-chart, the second stage begins with the resource allocation server calculating the number of data slots available for prioritized VBDC grant, for normal VBDC grant and for FCA grant. The resource allocation server then determines the number of data slots to be allocated to an access device through prioritized VBDC grant, which is the minimum of the access device's prioritized VBDC value (Dev.VBDC.PrioValue) and the number of data slots available for prioritized VBDC grant (AvailPrioVBDCCap). The first access device to whom prioritized VBDC grant cannot be given in whole or partially is tagged (LastPrioVBDCDev = Dev)

The resource allocation server then calculates the total number of credits available to the access device for prioritized VBDC grant. The total number of credits available is the minimum of the current total number of credits less the number of credits already used for prioritized VBDC grant (Dev.VBDC.TotalCredit + Dev.VBDC.Credit – Dev.VBDC.PrioValue), and the maximum number of credits the access device may accumulate for the subsequent frame (Dev.VBDC.MaxCredit – Dev.VBDC.Credit). The resource allocation server also updates the total number of data slots the access device has

2 An access device typically cannot send more data than the number of data slots per frequency

requested through VBDC (Dev.VBDC.ReqValue = Dev.VBDC.ReqValue – Dev.VBDC.PrioValue) and the total number of data slots available for prioritized VBDC grant (AvailPrioVBDCCap = AvailPrioVBDCCap – Dev.VBDC.PrioValue).

The resource allocation server continues to calculate the number of data slots to be allocated to an access device through normal VBDC grant, which is the minimum of the number of data slots available for normal VBDC grant (AvailVBDCCap), the number of data slots per frequency channel less the number of data slots to be allocated through CRA, RBDC and prioritized VBDC (ChannelCap – Dev.CRA.Value – Dev.RBDC.Value – Dev.VBDC.PrioValue) and the total number of data slots the access device has requested through VBDC (Dev.VBDC. ReqValue). The resource allocation server then updates the total number of data slots the access device has requested through VBDC (Dev.VBDC.ReqValue = Dev.VBDC.ReqValue – Dev.VBDC. Value) and the total number of data slots available for normal VBDC grant (AvailVBDCCap = AvailVBDCCap – Dev.VBDC.Value) as well as for prioritized VBDC grant (AvailPrioVBDCCap = AvailPrioVBDCCap – Dev.VBDC.Value). Similarly, the first access device to whom normal VBDC grant cannot be given in whole or partially is tagged (LastVBDC Dev = Dev).

At this point, the resource allocation server is able to determine the number of data slots it will grant to the access device through CRA, RBDC and VBDC (Dev.BandwidthAssign = Dev.CRA.Value + Dev.RBDC.Value + Dev.VBDC.PrioValue + Dev.VBDC.Value). The resource allocation server is free to distribute any unassigned bandwidth to access devices. The method that is used to distribute unassigned bandwidth is briefly described earlier and is shown in more detailed in Figure 3. The resource allocation server may grant a number of data slots (Dev.FCA.Value) up to the access device's accumulated FCA credits if the access device's bandwidth assignment through CRA, RBDC and VBDC is less than or equal to its FCA threshold, subject to the availability of data slots for FCA (AvailFCACap). The number of data slots granted through FCA will also be bounded by the number of data slots per frequency channel less the number of data slots to be allocated to the access device (ChannelCap – Dev.BandwidthAssign). Furthermore, if FCA is granted, the resource allocation server then updates the number of data slots available for prioritized VBDC grant, for normal VBDC grant and for FCA grant. On the other hand, the

channel.

first access device to whom FCA is not granted is tagged (LastFCADev = Dev).

The resource allocation server then calculates the total number of credits available to the access device for FCA grant. The total number of credits available is the minimum of the current total number of credits less the number of credits already used for FCA (Dev.FCA.TotalCredit + Dev.FCA.Credit – Dev.FCA.Value), and the maximum number of credits the access device may accumulate for the subsequent frame (Dev.FCA.MaxCredit – Dev.FCA.Credit).

At the end of the second stage, the resource allocation server updates the first entry of the table. The first entry of the table is updated to the first access device to whom prioritized VBDC grant cannot be given in whole or partially (LastPrioVBDC Dev), if any, or to the first access device to whom normal VBDC grant cannot be given in whole or partially (LastVBDC Dev), if any, or to the first access device to whom FCA grant cannot not granted (LastFCADev), if any, or to the entry directly after the current first entry in the table (Table.FirstEntry = Table.FirstEntry.Next).

An example of high-level software architecture for the implementation of the resource allocation method at the base station is illustrated in Figure 5.

The Transmission Controller component shown in Figure 5 is responsible for supplying data to the modem for transmission to the wireless network. Data for transmission to the wireless network can originate from the Wide Area Network (WAN), received through a set of Base Station Interfaces, as well as from the MAC Controller component in the form of MAC Protocol Data Unit (PDU). The Transmission Controller component is also responsible for receiving data from the modem. The component then determines if the received data is to be forwarded to the WAN through the Base Station Interfaces or to the MAC Controller component.

The MAC Controller component handles the MAC layer protocol. This component controls the flow of messages between the Base Station and the Access Nodes. It also takes care of internal items, including connection control and link provisioning and monitoring. This component receives information from the Client Characterization component for such purposes as client synchronization and power control. Relevant to the invention is the forwarding of bandwidth requests by the MAC Controller component to the Resource Allocation Server component. As described earlier, the resource allocation method distributes transmission bandwidth based on reservations and/or requests from each access device.

The Resource Allocation Server component is responsible for the task of distributing transmission bandwidth to each access device. This component uses the resource allocation method to perform this task. As described earlier, this component builds assignment tables (or BTPs) that regulate the transmission of each access device. These BTPs are forwarded to the MAC Controller component and are ultimately transmitted to all access devices.

Performance evaluations are conducted through discrete-event simulations. The simulation network model is composed of one hub node and 128 wireless access nodes, similar to a network model shown in Figure 6. The hub transmits data on the downlink channel, represented by a bus link. Each of the wireless access nodes is connected to the downlink bus so that each can receive all data sent by the hub. Each of the wireless access nodes is also connected to the uplink channel(s), represented also by a bus link. The hub is also connected to the uplink bus so that it can receive all data generated by each wireless access device.

The purpose of these simulations is to compare the performance of the invention with two other resource allocation methods described in [8] and [10]. Upstream high-priority and low-priority data traffic is generated in these simulations. High-priority data traffic is represented by the On-Off packet generator and is handled using the VBDC strategy. Low-priority data traffic is represented by the WWW packet generator as described in [11] and is also handled using the VBDC strategy. High-priority data traffic is set to receive up to 75% of capacity assigned to a wireless access device, while low-priority data traffic is set to receive the remaining capacity assigned to the access device. Similarly, each access device is eligible for FCA. Furthermore, each access device in the network accepts upstream traffic from 1 high-priority data source, for an average load of approximately 6 kbps per access device. The loading of low-priority data traffic in each access device is varied from an average load of approximately 6.5 kbps to 23.3 kbps.

The performance of four methods is compared. The first method is the resource allocation method described in [8], only an access device's VBDC requests never time-out. In this scheme, the resource allocation server maintains each VBDC request message sent by each wireless access device, and considers each access device's oldest VBDC request that has not been granted fully in distributing uplink capacity to all access devices. The second method is the resource allocation method described in [10]. This method is similar to the first method in that the resource allocation server maintains each VBDC request message sent by each access device.

The difference between the second method and the first method is that in the second method an access device's VBDC request that cannot be satisfied in full will be added to the access device's next VBDC request. The third method is the novel resource allocation method, only each access device's prioritized VBDC credit is set to zero. Thus, in the third method the resource allocation server maintains the total pending VBDC request from each access device. Newly arriving VBDC request from an access device will be added to the access device's total pending VBDC request. The fourth method is also the novel resource allocation method, with each access device's prioritized VBDC credit set to 1.0 and each access device's prioritized VBDC maximum total credit set to 4.0. Thus, each access device may have some prioritized VBDC request in each frame.

The frame structure in both simulations is similar. Each frame is approximately 24 ms in duration, consisting of two frequency channels with 16 synchronization slots and 128 traffic slots per channel. The effective data rate for each channel is thus 2 Mbps.

The propagation delay between the hub and the closest wireless access device is approximately 125 msec, while the delay between the hub and the furthest access device is approximately 130 msec.

Each simulation is conducted for a simulated time of 350 seconds. Simulation statistics, however, are collected starting from a simulated time of 50 seconds.

Figure 7 shows the average packet transfer delay of high-priority data traffic for the four methods. As can be seen from Figure 7, the average delay experienced by high-priority data traffic is similar for the four methods up to a total traffic load of approximately 80%. At a total traffic load of approximately 92%, however, the fourth method, which is the new resource allocation method with prioritized VBDC credit associated to each access device, performs significantly better than the second or the third method. This is because at such a high load, the total capacity requested by each access device becomes very large at peak periods, resulting in only a few access devices capable of transmitting their high-priority data traffic in each frame when the second or the third method is employed. When the fourth method is employed, however, each access device often has prioritized VBDC capacity at each frame, so that most access devices are capable of transmitting their high-priority data traffic in each frame, thus reducing the average packet transfer delay significantly. The first method is also capable of reducing the average packet transfer delay of high-priority data traffic, although not as effectively

as the fourth method, due to its strategy of not accumulating VBDC capacity requests from each access device. The first method, in effect, allows more access devices to transmit their traffic in each frame than the second or the third method does.

Fig. 8 shows the average packet transfer delay of low-priority data traffic for the four methods. As shown in the figure, the average delay of low-priority data traffic is maintained low up to a total traffic load of approximately 80%. The average delay increases dramatically as the traffic load increases above 80%. This is due to the burstiness of WWW traffic. An important point to note from the figure is that the average delay of low-priority data traffic achieved by the fourth method is very close to that achieved by the second or the third method, and is lower than that achieved by the first method. The significant delay performance enhancement offered by the fourth method to high-priority data traffic results in little degradation in delay performance offered to low-priority data traffic. Furthermore, the fourth method outperforms the first method in both fronts; it achieves lower average packet transfer delay values at very high traffic loads for both high-priority and low-priority data traffic.

The main advantages of the present invention are as follows:

- It allows the network to make use of the VBDC method to transport non-real-time traffic with bandwidth requirements in order to increase bandwidth utilization and reduce data transfer delay of such traffic.
- It allows the network to provide prioritized access to certain non-real-time traffic when utilizing the VBDC method.
- It allows the network to provide fair access to non-real-time traffic using the VBDC method.
- It allows the network to shape non-real-time traffic, reducing its short-term burstiness and congestion loss.

GLOSSARY OF ACRONYMS

ATM	Asynchronous Transfer Mode
BoD	Bandwidth on Demand
BTP	Burst Time Plan
CDMA	Code Division Multiple Access
CF-DAMA	Combined Free – Demand Assignment Multiple Access
CRA	Continuous Rate Allocation
FCA	Free Capacity Allocation
FDMA	Frequency Division Multiple Access
MAC	Media Access Control
MF-TDMA	Multi-Frequency Time Division Multiple Access
PDU	Protocol Data Unit
QoS	Quality of Service
RBDC	Rate-Based Dynamic Capacity
TDMA	Time Division Multiple Access
VBDC	Volume-Based Dynamic Capacity
WAN	Wide Area Network

We claim:

1. A method for resource allocation applicable to a base station unit in broadband terrestrial wireless and satellite networks, comprising the steps of:

(i) performing a first scan of an access device table by calculating the number of data slots for prioritized VBDC grant and for normal VBDC grant to each wireless access device, and calculating the total number of data slots for prioritized VBDC grant and for normal VBDC grant to all wireless access devices in the access device table.

(ii) performing a second scan of the access device table by calculating the total number of data slots available for prioritized VBDC grant, for normal VBDC grant and for FCA grant, calculating the number of data slots to be allocated to each access device through prioritized VBDC grant, calculating the total number of credits available to each access device for prioritized VBDC grant, calculating the number of data slots to be allocated to each access device through normal VBDC grant, calculating the total number of data slots to be allocated to each access device, which is the sum of data slots reserved through CRA and RBDC as well as data slots obtained through prioritized and normal VBDC grants, determining if each access device is eligible to receive FCA, calculating the number of data slots to be allocated to each access device through FCA grant, and calculating the total number of credits available to each access device for FCA grant.

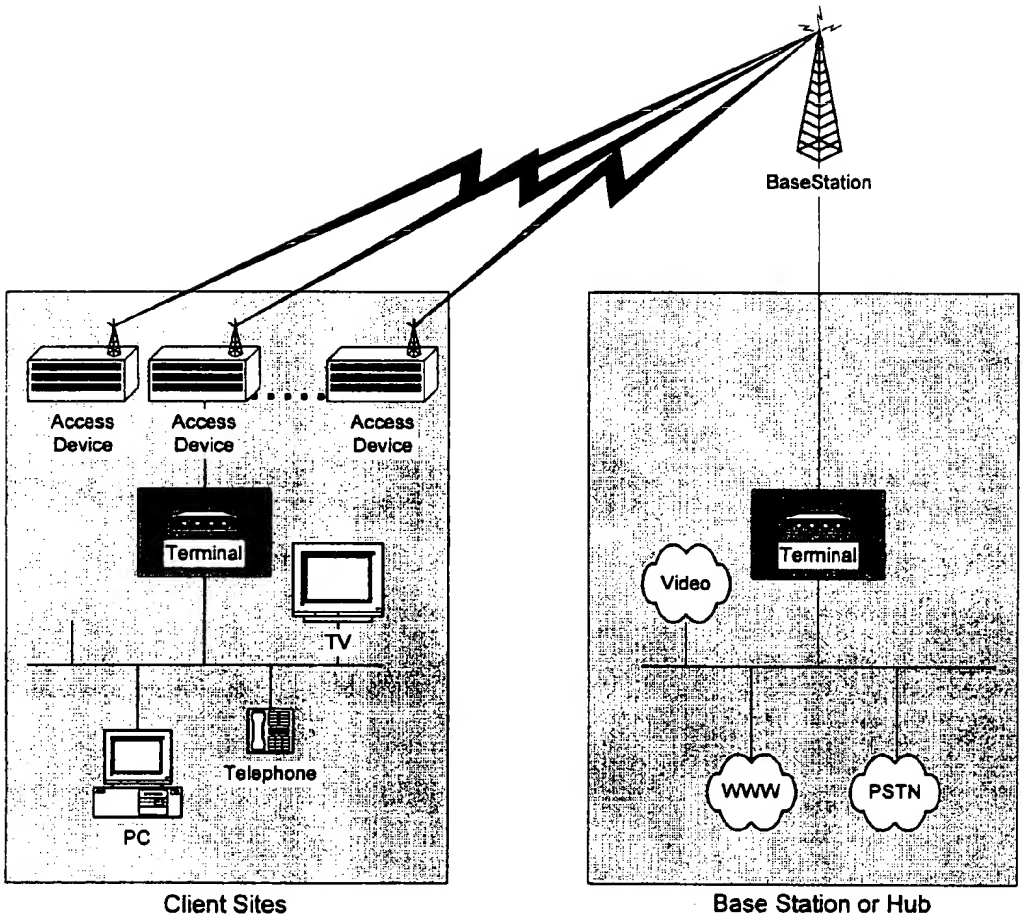
BaseStation
Access Device
Access Device
Access Device
Terminal
TV
Telephone
PC
Client Sites
Terminal
Video
WWW
PSTN
Base Station or Hub

Figure 1

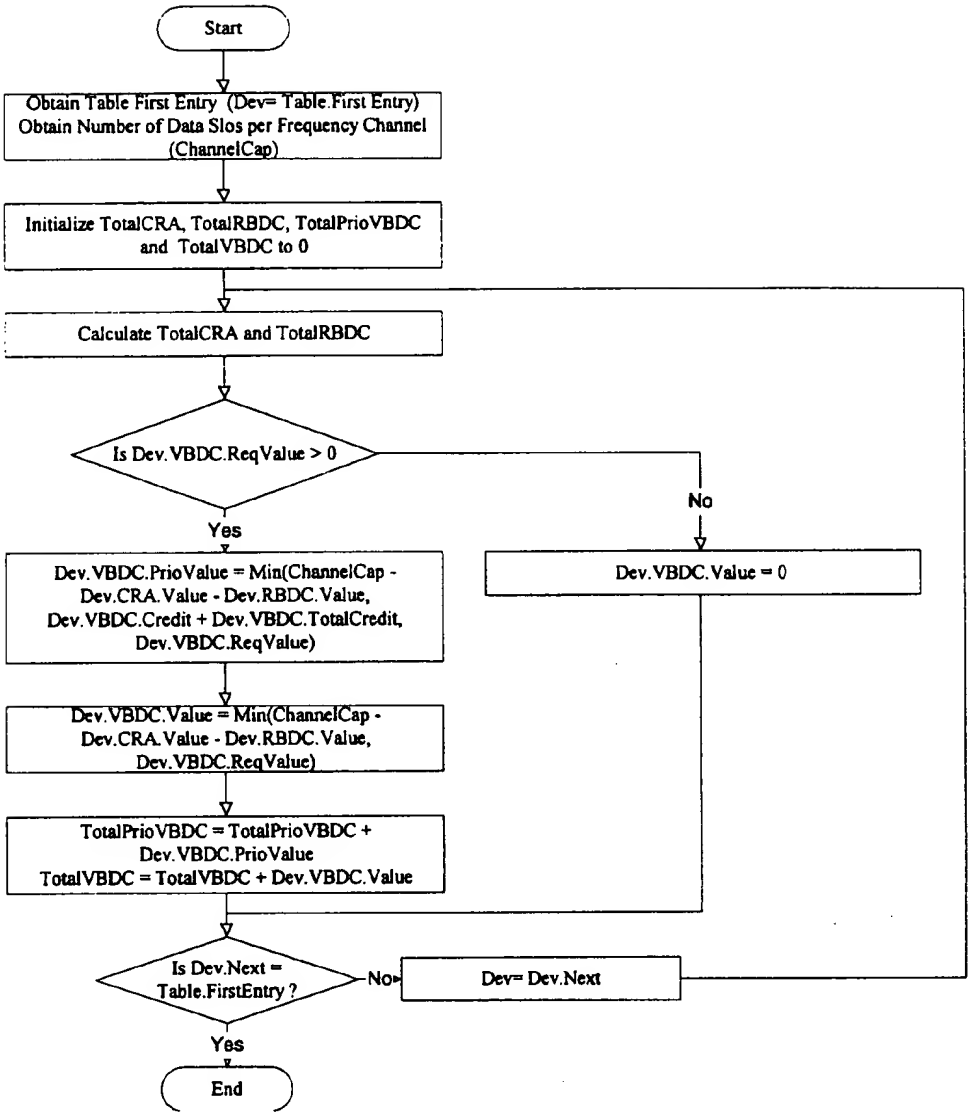
Start
Obtain Table First Entry (Dev= Table.First Entry)
Obtain Number of Data Slos per Frequency Channel (ChannelCap)
Initialize TotalCRA, TotalRBDC, TotalPrioVBDC and TotalVBDC to 0
Calculate TotalCRA and TotalRBDC
Is Dev.VBDC.ReqValue > 0
No
Yes
Dev.VBDC.PrioValue = Min(ChannelCap - Dev.CRA.Value - Dev.RBDC.Value, Dev.VBDC.Credit + Dev.VBDC.TotalCredit, Dev.VBDC.ReqValue)
Dev.VBDC.Value = 0
Dev.VBDC.Value = Min(ChannelCap - Dev.CRA.Value - Dev.RBDC.Value, Dev.VBDC.ReqValue)
TotalPrioVBDC = TotalPrioVBDC + Dev.VBDC.PrioValue
TotalVBDC = TotalVBDC + Dev.VBDC.Value
Is Dev.Next = Table.FirstEntry ?
No
Dev= Dev.Next
Yes
End

Figure 2

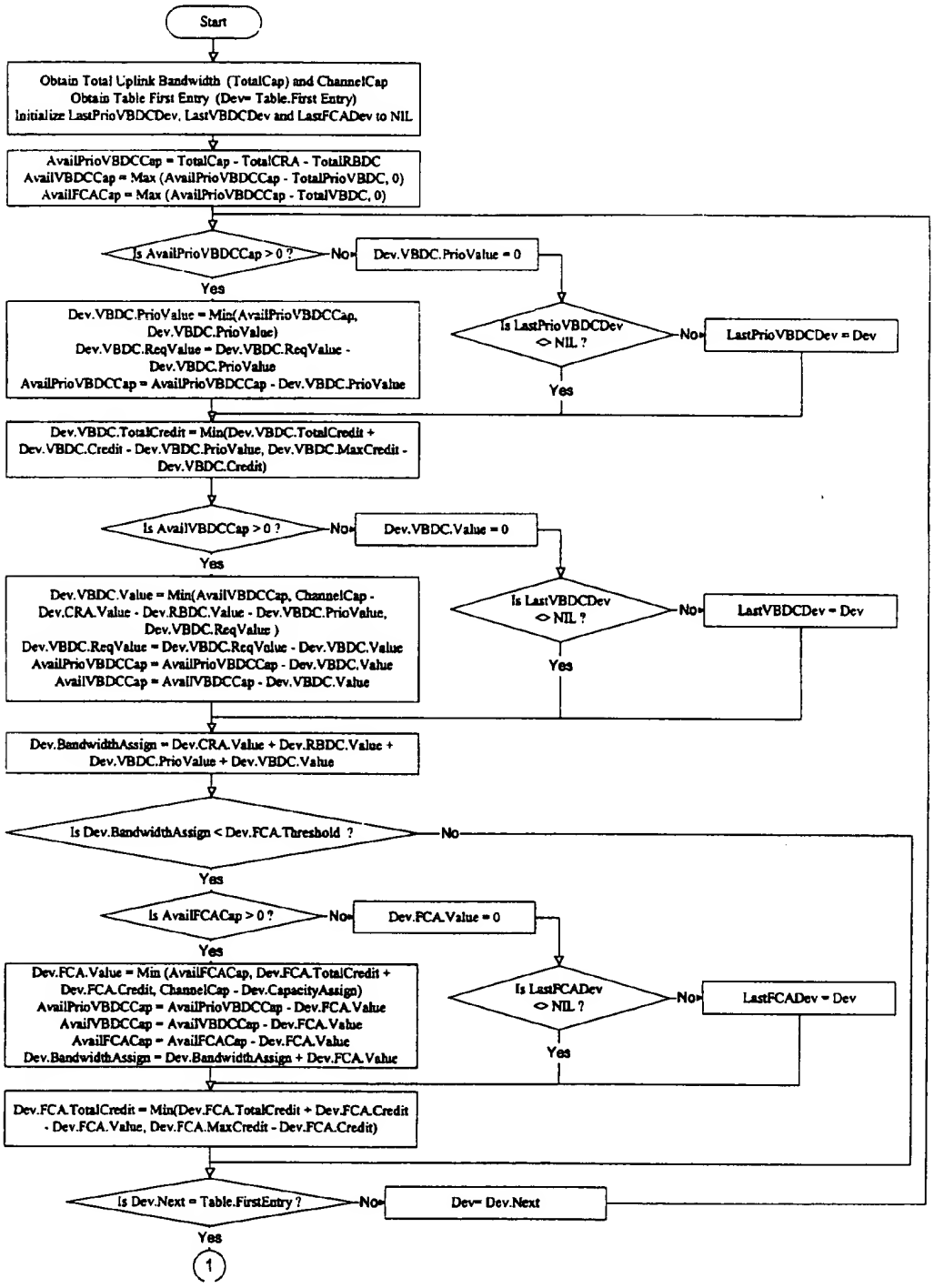
Start
Obtain Total Uplink Bandwidth (TotalCap) and ChannelCap
Obtain Table First Entry (Dev= Table.First Entry)
Initialize LastPrioVBDCDev, LastVBDCDev and LastFCADev to NIL
AvailPrioVBDCCap = TotalCap - TotalCRA - TotalRBDC
AvailVBDCCap = Max (AvailPrioVBDCCap - TotalPrioVBDC, 0)
AvailFCACap = Max (AvailPrioVBDCCap - TotalVBDC, 0)
Is AvailPrioVBDCCap > 0 ?
No
Dev.VBDC.PrioValue = 0
Yes
Dev.VBDC.PrioValue = Min(AvailPrioVBDCCap, Dev.VBDC.PrioValue)
Dev.VBDC.ReqValue = Dev.VBDC.ReqValue - Dev.VBDC.PrioValue
AvailPrioVBDCCap = AvailPrioVBDCCap - Dev.VBDC.PrioValue
Is LastPrioVBDCDev <> NIL ?
No
LastPrioVBDCDev = Dev
Yes
Dev.VBDC.TotalCredit = Min(Dev.VBDC.TotalCredit + Dev.VBDC.Credit - Dev.VBDC.PrioValue, Dev.VBDC.MaxCredit - Dev.VBDC.Credit)
Is AvailVBDCCap > 0 ?
No
Dev.VBDC.Value = 0
Yes
Dev.VBDC.Value = Min(AvailVBDCCap, ChannelCap - Dev.CRA.Value - Dev.RBDC.Value - Dev.VBDC.PrioValue, Dev.VBDC.ReqValue)
Dev.VBDC.ReqValue = Dev.VBDC.ReqValue - Dev.VBDC.Value
AvailPrioVBDCCap = AvailPrioVBDCCap - Dev.VBDC.Value
AvailVBDCCap = AvailVBDCCap - Dev.VBDC.Value
Is LastVBDCDev <> NIL ?
No
LastVBDCDev = Dev
Yes
Dev.BandwidthAssign = Dev.CRA.Value + Dev.RBDC.Value + Dev.VBDC.PrioValue + Dev.VBDC.Value
Is Dev.BandwidthAssign < Dev.FCA.Threshold ?
No
Yes
Is AvailFCACap > 0 ?
No
Dev.FCA.Value = 0
Yes
Dev.FCA.Value = Min (AvailFCACap, Dev.FCA.TotalCredit + Dev.FCA.Credit, ChannelCap - Dev.CapacityAssign)
AvailPrioVBDCCap = AvailPrioVBDCCap - Dev.FCA.Value
AvailVBDCCap = AvailVBDCCap - Dev.FCA.Value
AvailFCACap = AvailFCACap - Dev.FCA.Value
Dev.BandwidthAssign = Dev.BandwidthAssign + Dev.FCA.Value
Is LastFCADev <> NIL ?
No
LastFCADev = Dev
Yes
Dev.FCA.TotalCredit = Min(Dev.FCA.TotalCredit + Dev.FCA.Credit - Dev.FCA.Value, Dev.FCA.MaxCredit - Dev.FCA.Credit)
Is Dev.Next = Table.FirstEntry ?
No
Dev= Dev.Next
Yes
1

Figure 3

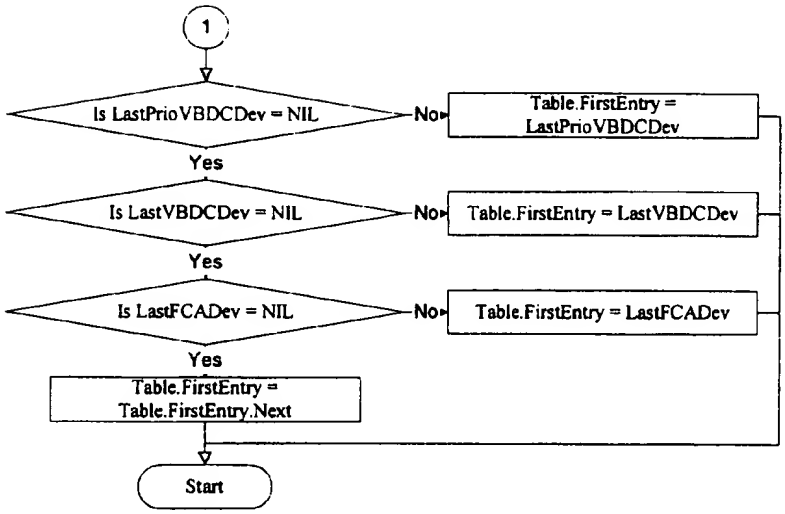

Figure 4

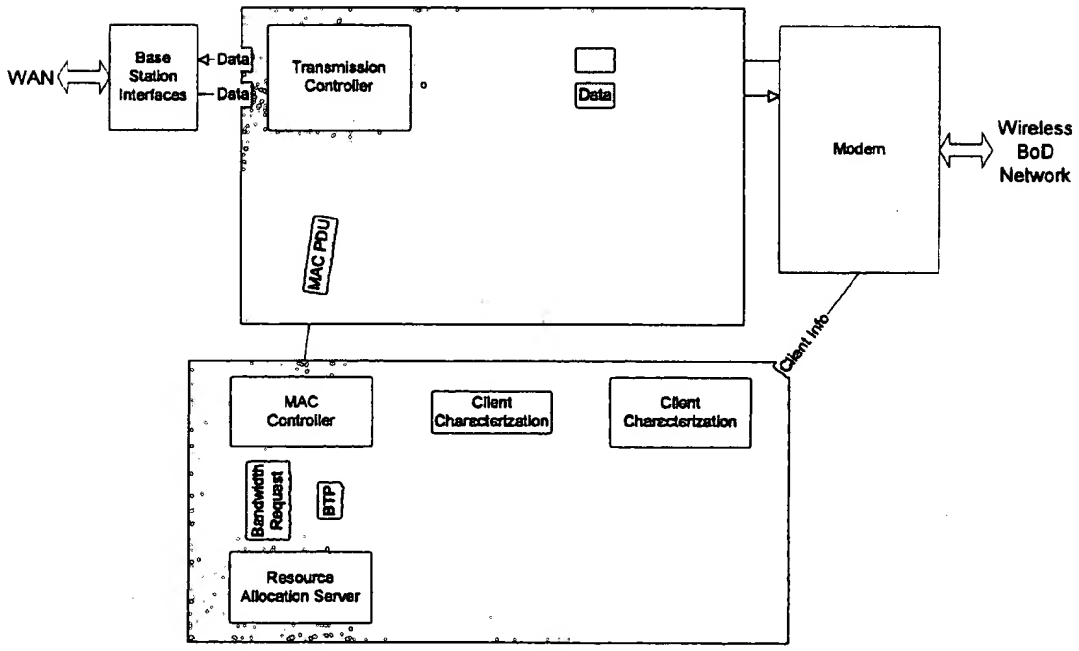

Figure 5

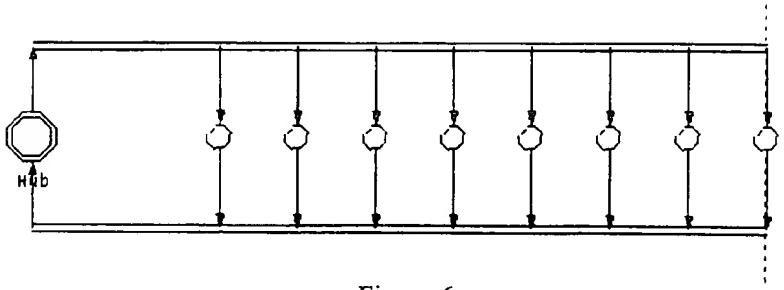

Figure 6

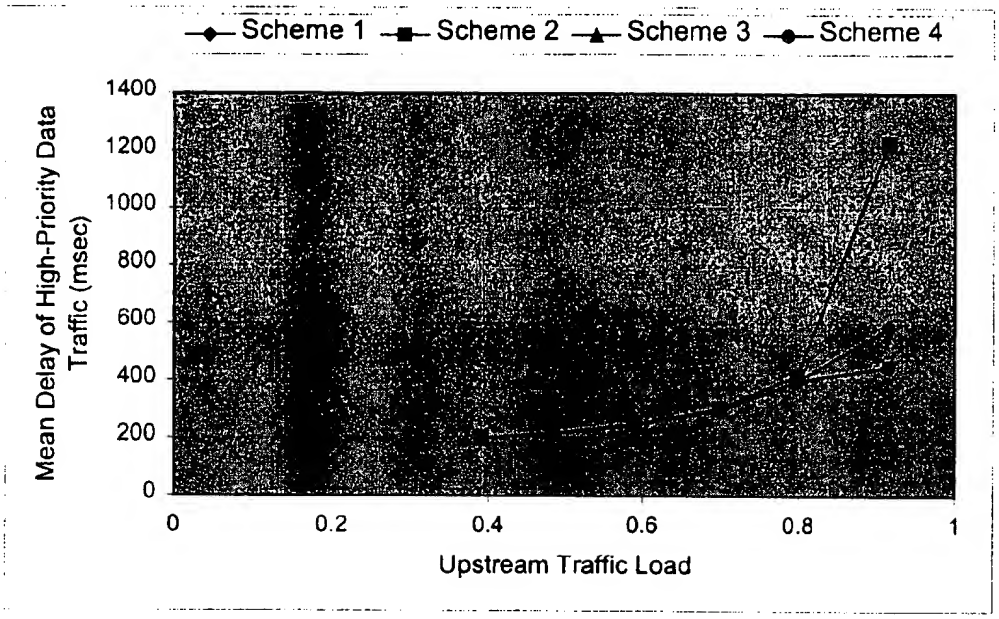

Figure 7

resource

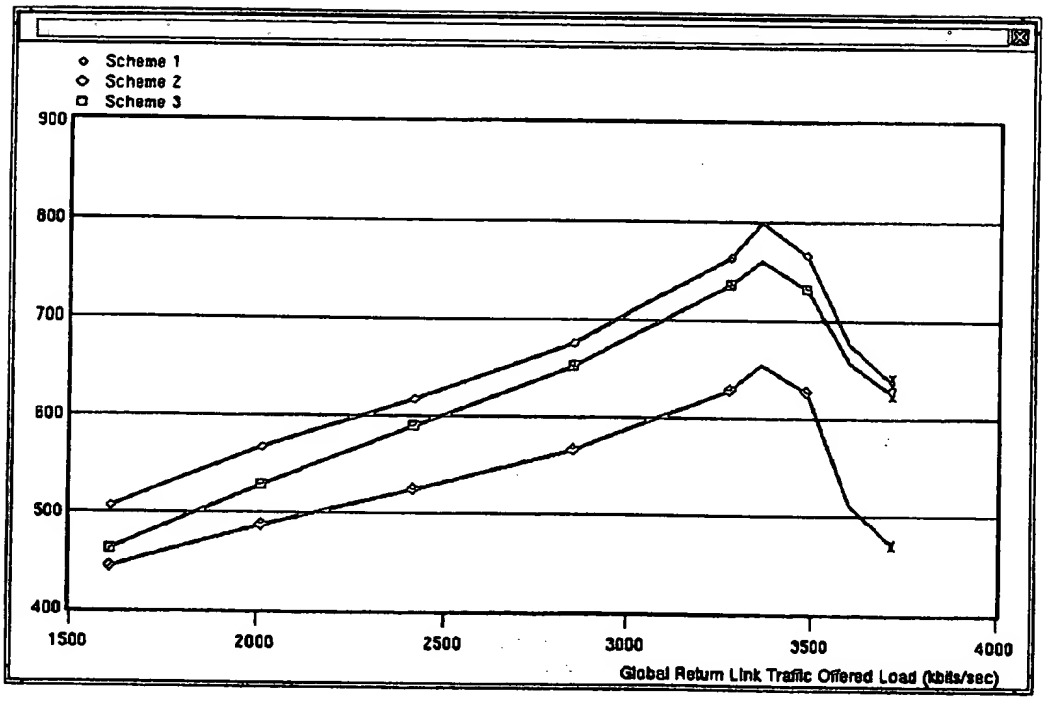

Figure 8